



Bölüm 4

Örneklem Dağılımları



UZAKTAN EĞİTİM UYGULAMA VE ARAŞTIRMA MERKEZİ

telefon 0(312) 202 82 00 • eposta guzem@gazi.edu.tr • adres Gazi Üniversitesi Rektörlük Binası No:6/f

guzem.gazi.edu.tr • uzaktanegitim.gazi.edu.tr • lms.gazi.edu.tr

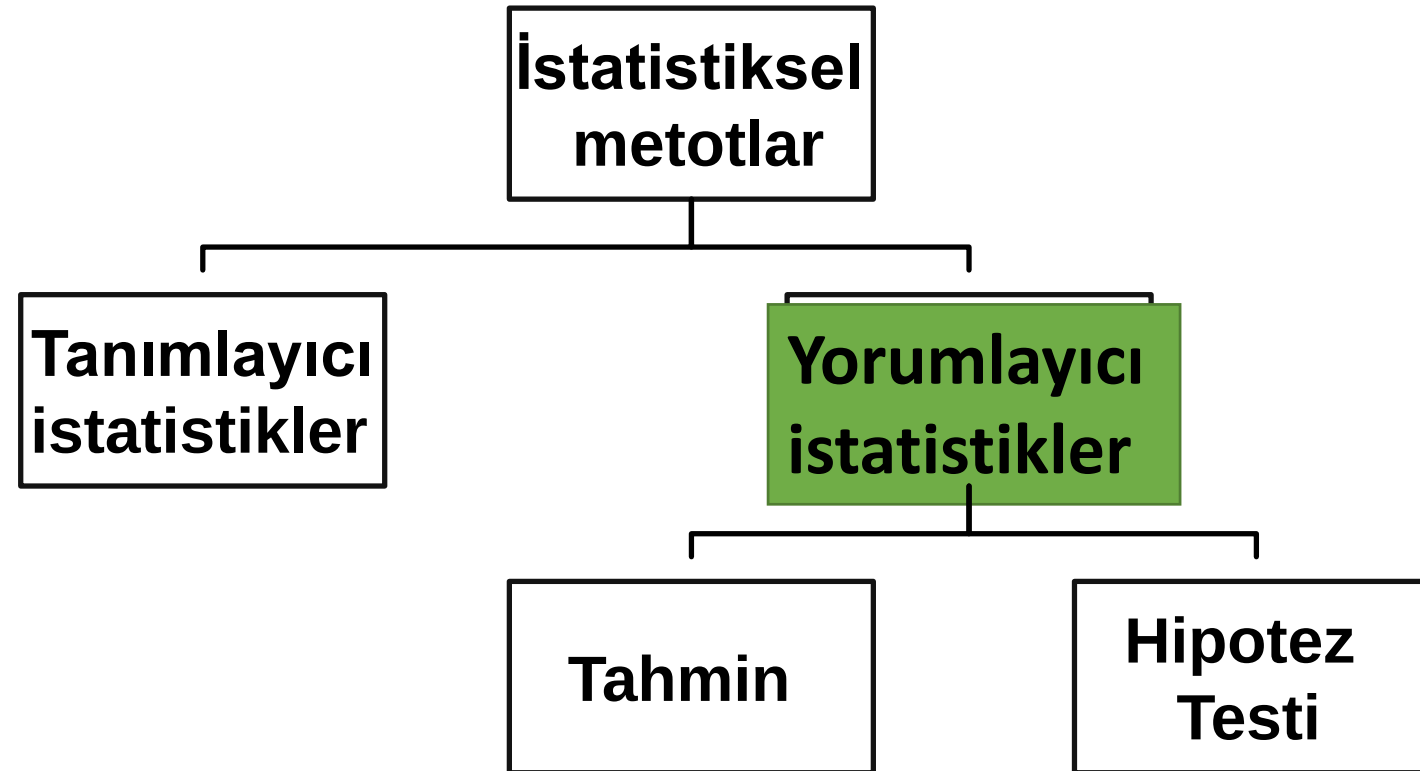
Öğrenme hedefleri

Bu bölümde:

- Örneklem dağılımı kavramını,
- Örnek ortalaması ve örnek oranlarıyla bağlantılı olarak olasılıkların hesaplanmasını,
- Merkezi Limit Teoreminin önemini,

inceleyeceğiz

İstatistiksel metotlar

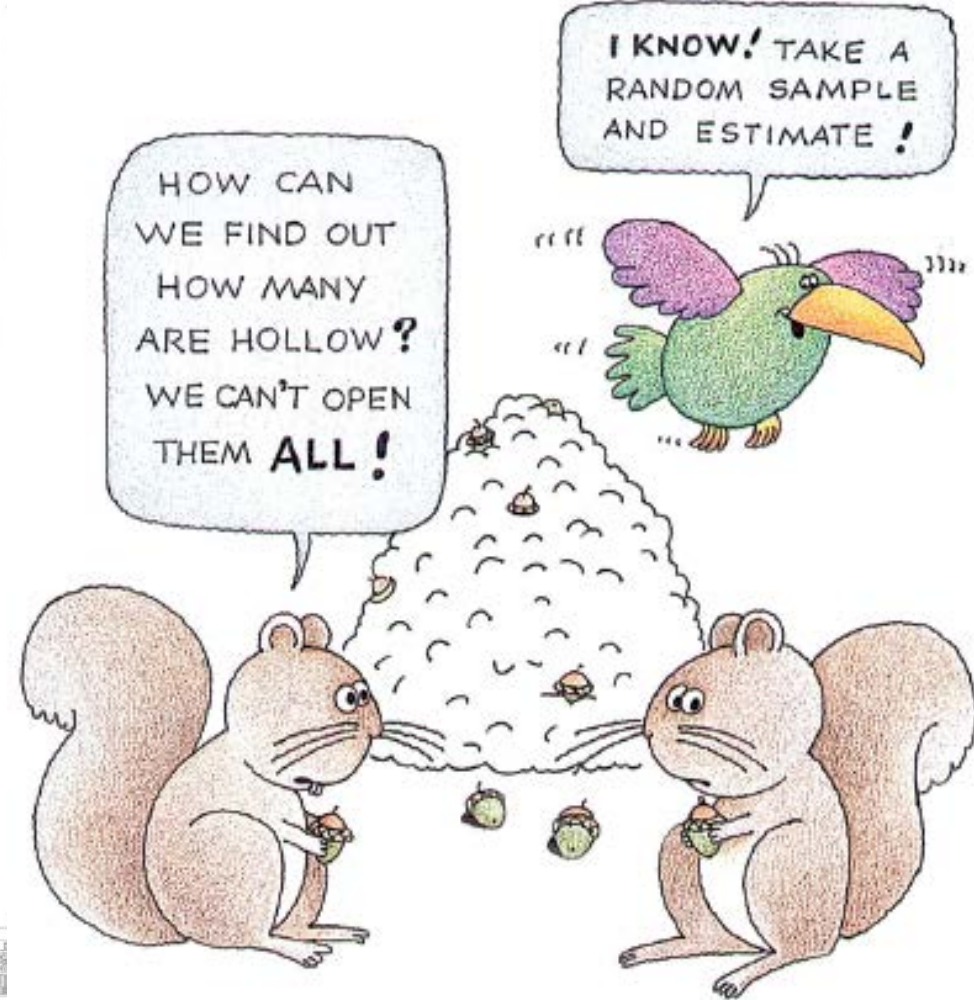
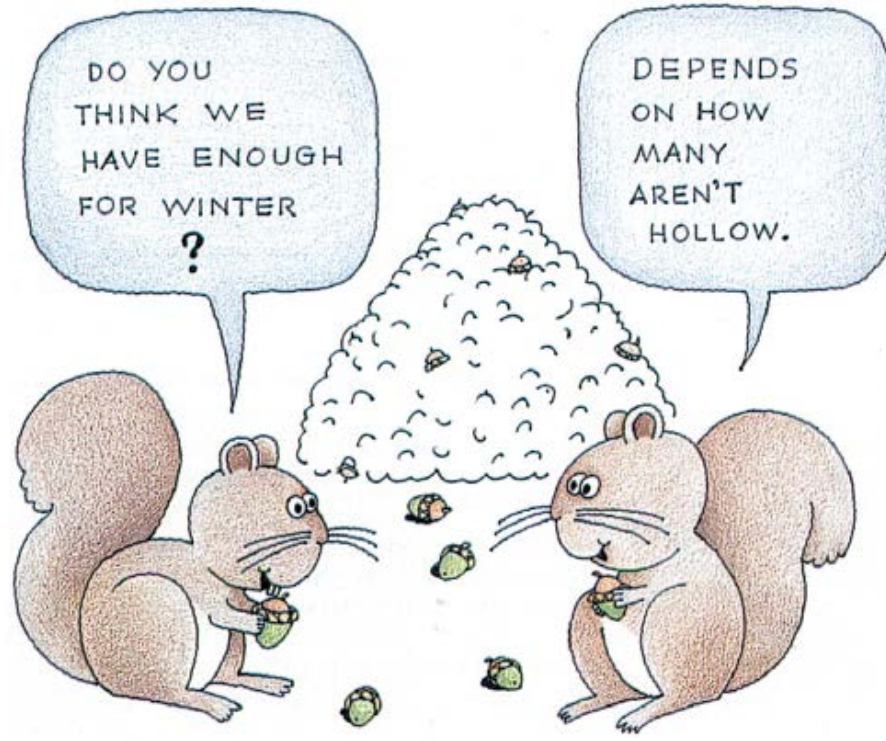


Yorumlayıcı istatistikler

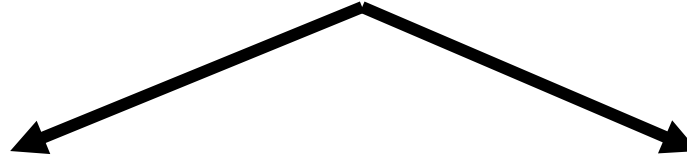
- Aralık tahminleme ve hipotez testlerini içerir.
- Amacı populasyon karakteristikleri hakkında karar vermektir.

Populasyon?





İstatistiksel Tahminleme



Nokta Tahmini

Populasyon parametresinin tek bir tahmin değerini verir

$$\bar{X} = \hat{\mu}$$

$$s = \hat{\sigma}$$

$$p = \hat{P}$$

Aralık Tahmini

Populasyon parametresinin tahmin aralığını verir. Nokta tahmini kullanılarak hesaplanır.

$$20 \leq \mu \leq 60$$

$$2.5 \leq \sigma^2 \leq 3.4$$

$$0.25 \leq P \leq .035$$

Tanım



- **Tahmin (kestirim veya öngörü):**
- Mevcut bilgi ve deneylere dayanarak olayın bütünü hakkında bir yargıya varmaktır.
 - Bu anlamda, anakütleden çekilen örnek verileri üzerinde istatistiksel yöntemler uygulanarak elde edilen sonuçlar anakütleye genelleştirilir.

İstatistiksel çıkarsama

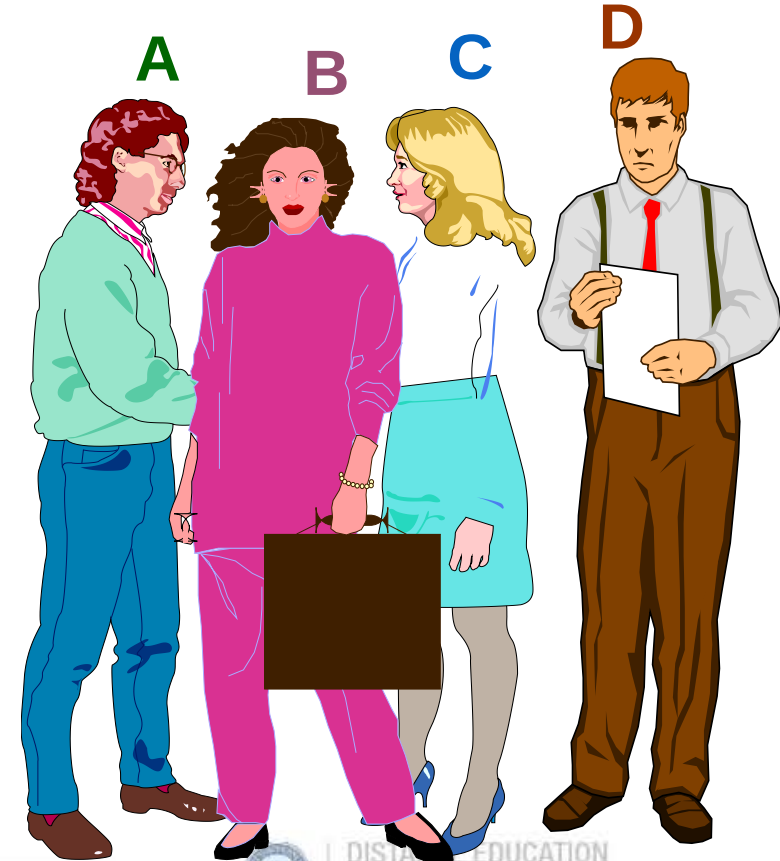
- İstatistiksel metotlar popölasyon hakkında karar vermek veya sonuç çıkarmak için kullanılır.
- İstatistiğin bu yönü genel olarak **istatistiksel çıkarsama** (statistical inference) olarak bilinir.
- Bu teknikler sonuç çıkarsamada örneklemdaki bilgiyi kullanırlar.
- Bunun için anakütleyi temsil eden örnekler çekilir.

Örneklem dağılımları

- Bir örneklem dağılımı, bir popülasyondan seçilen belirli boyuttaki bir örnek için bir örnek istatistiğinin tüm olası değerlerinin dağılımıdır.
- Örneğin, üniversitenizden ortalama GNO (genel not ortalaması) ile ilgili olarak 50 öğrenciyi örneklediğinizi varsayalım. Eğer farklı 50 kişilik numuneler aldıysanız, her numune için farklı bir ortalama hesaplayacaksınız. İşte biz de herhangi bir 50 öğrenci örneği için hesaplayabileceğimiz potansiyel ortalama GNO'ların dağılımı ile ilgileniyoruz.

Bir örneklem dağılımının geliştirilmesi

- Şöyle bir popülasyon olduğunu düşünelim...
- Popülasyon büyüklüğü $N=4$
- X rassal değişkeni,
bireylerin **yaşıdır**
- X değerleri: 18, 20,
22, 24
- Örneklem sayısı: $n=2$

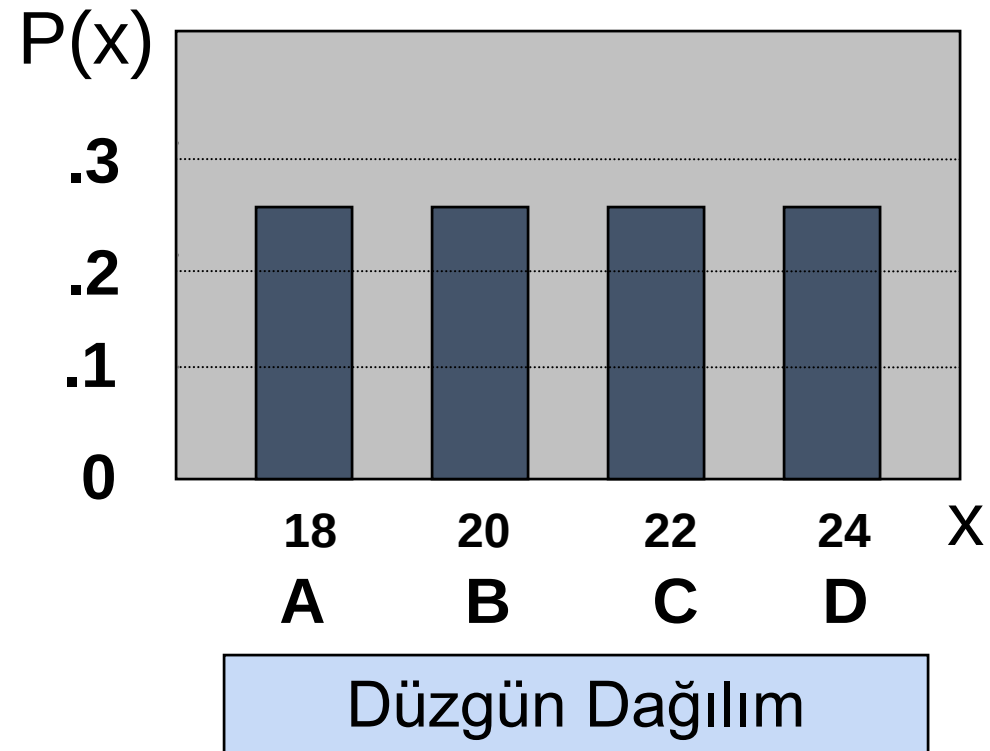


Bir örneklem dağılımının geliştirilmesi

Popülasyon Dağılımının Özet Ölçüleri:

$$\begin{aligned}\mu &= \frac{\sum x_i}{N} \\ &= \frac{18 + 20 + 22 + 24}{4} = 21\end{aligned}$$

$$\sigma = \sqrt{\frac{\sum (x_i - \mu)^2}{N}} = 2.236$$



Bir örneklem dağılımının geliştirilmesi

- $n=2$ büyüklüğüne sahip tüm olası örnekleri düşünelim

1. Gözlem	2. Gözlem			
	18	20	22	24
18	18,18	18,20	18,22	18,24
20	20,18	20,20	20,22	20,24
22	22,18	22,20	22,22	22,24
24	24,18	24,20	24,22	24,24

16 olası örnek
(yerdeğiştirme ile
örnekleme)



1. Gözlem	2. Gözlem			
	18	20	22	24
18	18	19	20	21
20	19	20	21	22
22	20	21	22	23
24	21	22	23	24

16 Örnek
Ortalaması

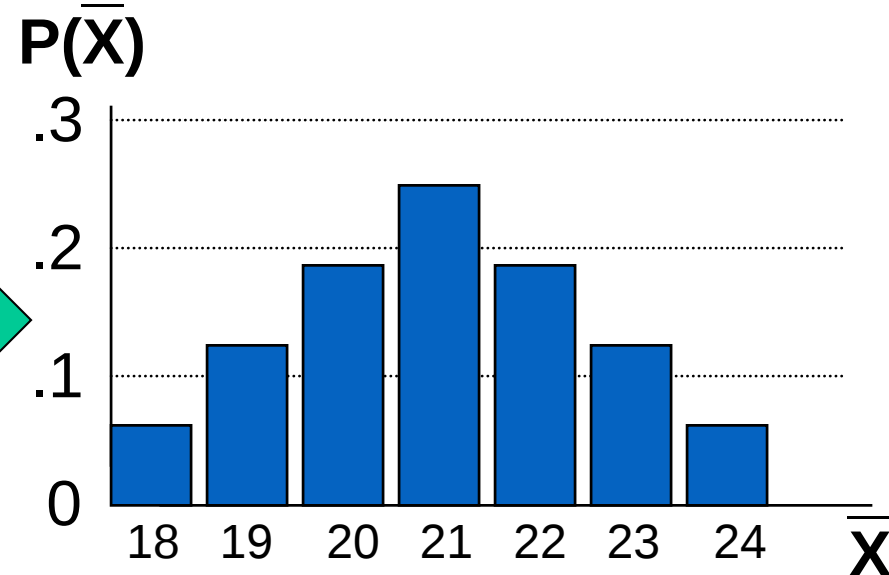
Bir örneklem dağılımının geliştirilmesi

- Tüm Örnek Ortalamalarının Örneklem Dağılımı

16 Örnek Ortalaması

1. Gözlem	2. Gözlem			
	18	20	22	24
18	18	19	20	21
20	19	20	21	22
22	20	21	22	23
24	21	22	23	24

Örnek ortalaması dağılımı



(artık düzgün dağılım değil)

Bir örneklem dağılımının geliştirilmesi

- Bu Örneklem Dağılımının Özet Ölçüleri:

$$\mu_{\bar{X}} = \frac{18 + 19 + 19 + \dots + 24}{16} = 21$$

$$\sigma_{\bar{X}} = \sqrt{\frac{(18 - 21)^2 + (19 - 21)^2 + \dots + (24 - 21)^2}{16}} = 1.58$$

Not: 2 büyüklüğüne sahip 16 farklı numune bulunduğu için burada 16 ile bölme yapılıyor.

Popülasyon dağılımı ile örnek ortalaması dağılımının kıyaslanması

Popülasyon

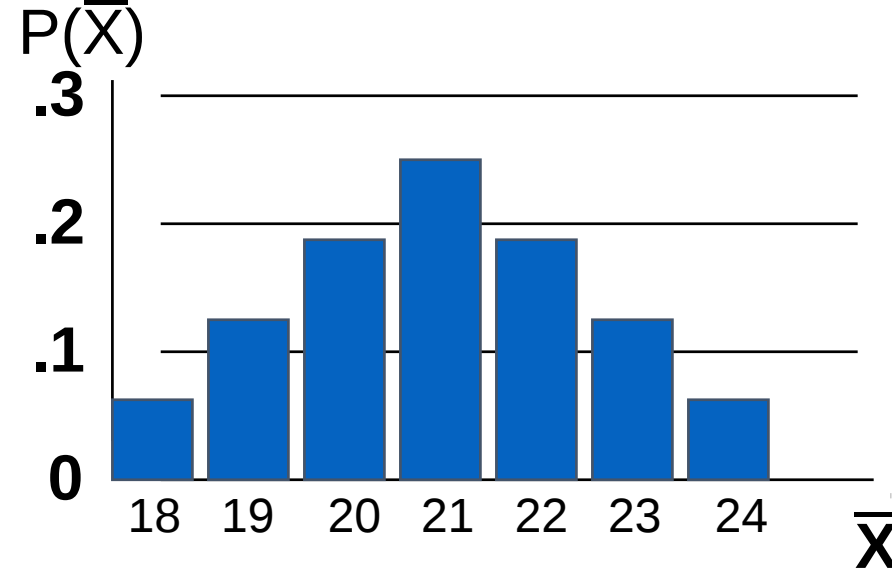
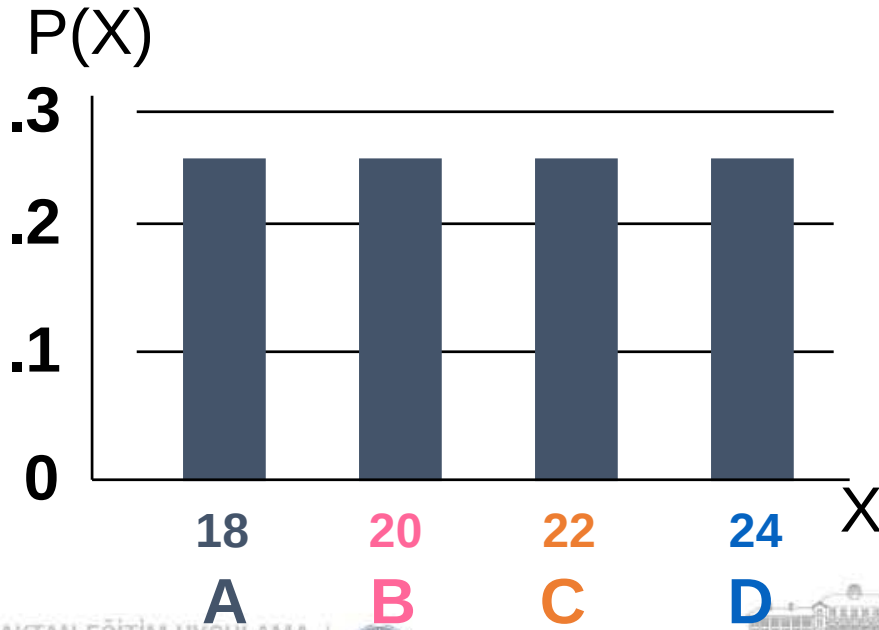
$N = 4$

$\mu = 21$ $\sigma = 2.236$

Örnek Ortalaması Dağılımı

$n = 2$

$\mu_{\bar{X}} = 21$ $\sigma_{\bar{X}} = 1.58$



Örnek ortalaması örneklem dağılımı: Ortalamanın standart hatası

- Aynı popülasyondan aynı boyuttaki farklı numuneler, farklı örnek ortalamaları sağlayacaktır
- Örnekten örneğe ortalamadaki değişkenliğin bir ölçüsü, **Ortalamanın Standart Hatası** ile verilir:

(Bu, örnekleme için sonsuz bir popülasyondan yer değiştirme yaparak ya da yer değiştirme yapmadan gerçekleştiğini varsaymaktadır)

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

- Örnekleme boyutu arttıkça ortalamanın standart hatası azalmaktadır

Örnek ortalaması örneklem dağılımı: Eğer popülasyon normal dağılmışsa

- Bir popülasyon, μ ortalaması ve σ standart sapması ile **normal** dağılmışsa, $\bar{X}$ 'in örneklem dağılımı da aşağıdaki parametrelerle **normal dağılıma uymaktadır**.

$$\mu_{\bar{X}} = \mu$$

ve

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$$

Ortalamanın örneklem dağılımı için Z-değeri

- $\bar{X}$ 'in örneklem dağılımı için Z-değeri :

$$Z = \frac{(\bar{X} - \mu_{\bar{X}})}{\sigma_{\bar{X}}} = \frac{(\bar{X} - \mu)}{\frac{\sigma}{\sqrt{n}}}$$

$\bar{X}$ = örnek ortalaması

μ = popülasyon ortalaması

σ = popülasyon standart sapması

n = örnek büyüklüğü

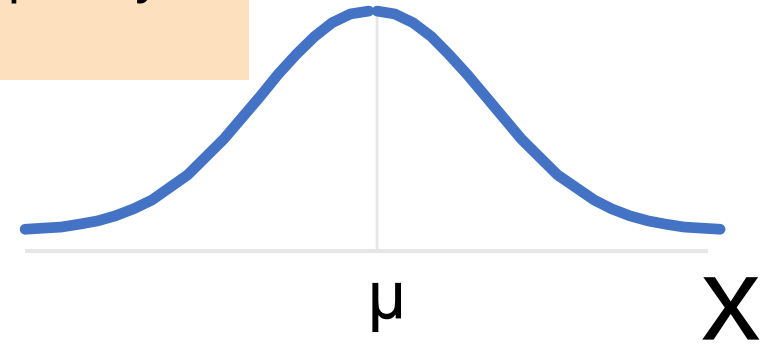
Örneklem dağılımı özellikleri

-

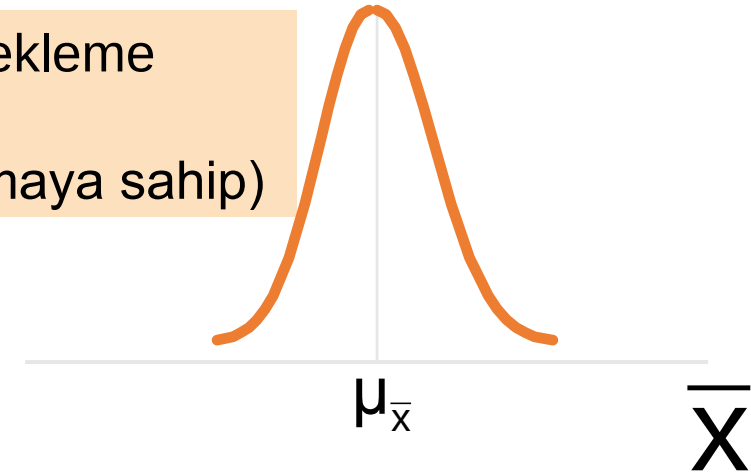
$$\mu_{\bar{X}} = \mu$$

(yani $\bar{X}$ yansızdır)

Normal Popülasyon
Dağılımı

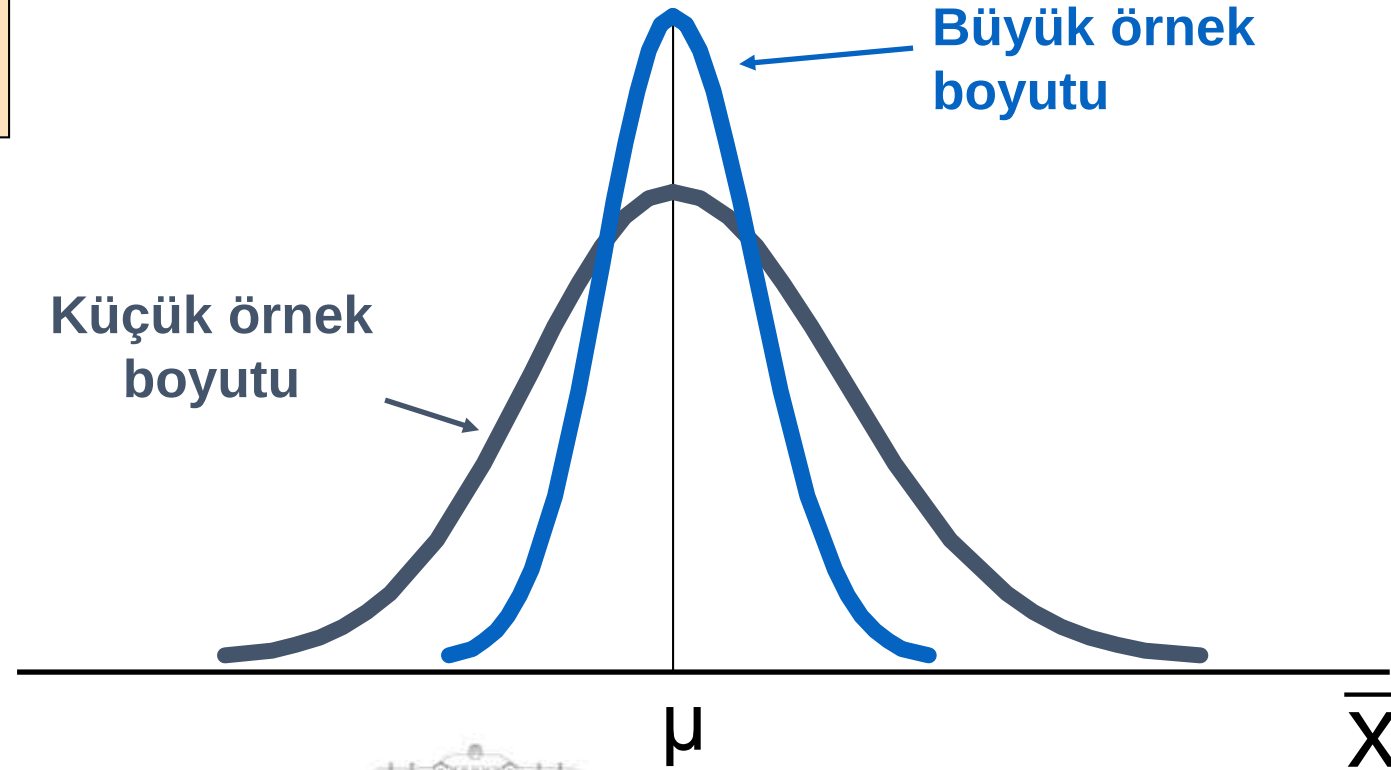


Normal Örneklem
Dağılımı
(aynı ortalamaya sahip)



Örneklem dağılımı özellikleri

n arttıkça,
 $\sigma_{\bar{x}}$ azalmaktadır



Örneklem ortalamasının sabit bir oranını içeren bir aralığın belirlenmesi

$\mu = 368$, $\sigma = 15$ ve $n = 25$ olduğunda örnek ortalamalarının % 95'ini içerecek μ çevresinde simetrik olarak dağılmış bir aralık bulun.

- o Aralık, örneklem ortalamalarının % 95'ini içerdiğinden, örneklem ortalamalarının % 5'i aralığın dışında olacak demektir
- o Aralık simetrik olduğu için % 2.5 üst sınırın üstünde olacak ve % 2.5 alt sınırın altına gelecektir.
- o Standart normal dağılım tablosundan, % 2.5 (0.0250) altında Z skorunun -1.96, ve % 2.5 (0.0250) üstünde Z skorunun 1.96 olduğu görülmektedir.

Örneklem ortalamasının sabit bir oranını içeren bir aralığın belirlenmesi

- Aralığın alt sınırının hesaplanması

$$\bar{X}_L = \mu + Z \frac{\sigma}{\sqrt{n}} = 368 + (-1.96) \frac{15}{\sqrt{25}} = 362.12$$

- Aralığın üst sınırının hesaplanması

$$\bar{X}_U = \mu + Z \frac{\sigma}{\sqrt{n}} = 368 + (1.96) \frac{15}{\sqrt{25}} = 373.88$$

- Örnek büyüklüğü 25 olan tüm örneklerin ortalamasının %95'i 362.12 ile 373.88 aralığındadır.

Örnek ortalaması örneklem dağılımı: popülasyon normal dağılıma uymadığında

- **Merkezi Limit Teoremini** uygulayabiliriz :
 - o Popülasyon **normal dağılıma uymasa** bile,
 - o Örneklem büyüklüğü yeterince büyük olduğu sürece popülasyondan elde edilen örnek ortalamaları **yaklaşık olarak normal** olacaktır.

Örnekleme dağılımının özellikleri:

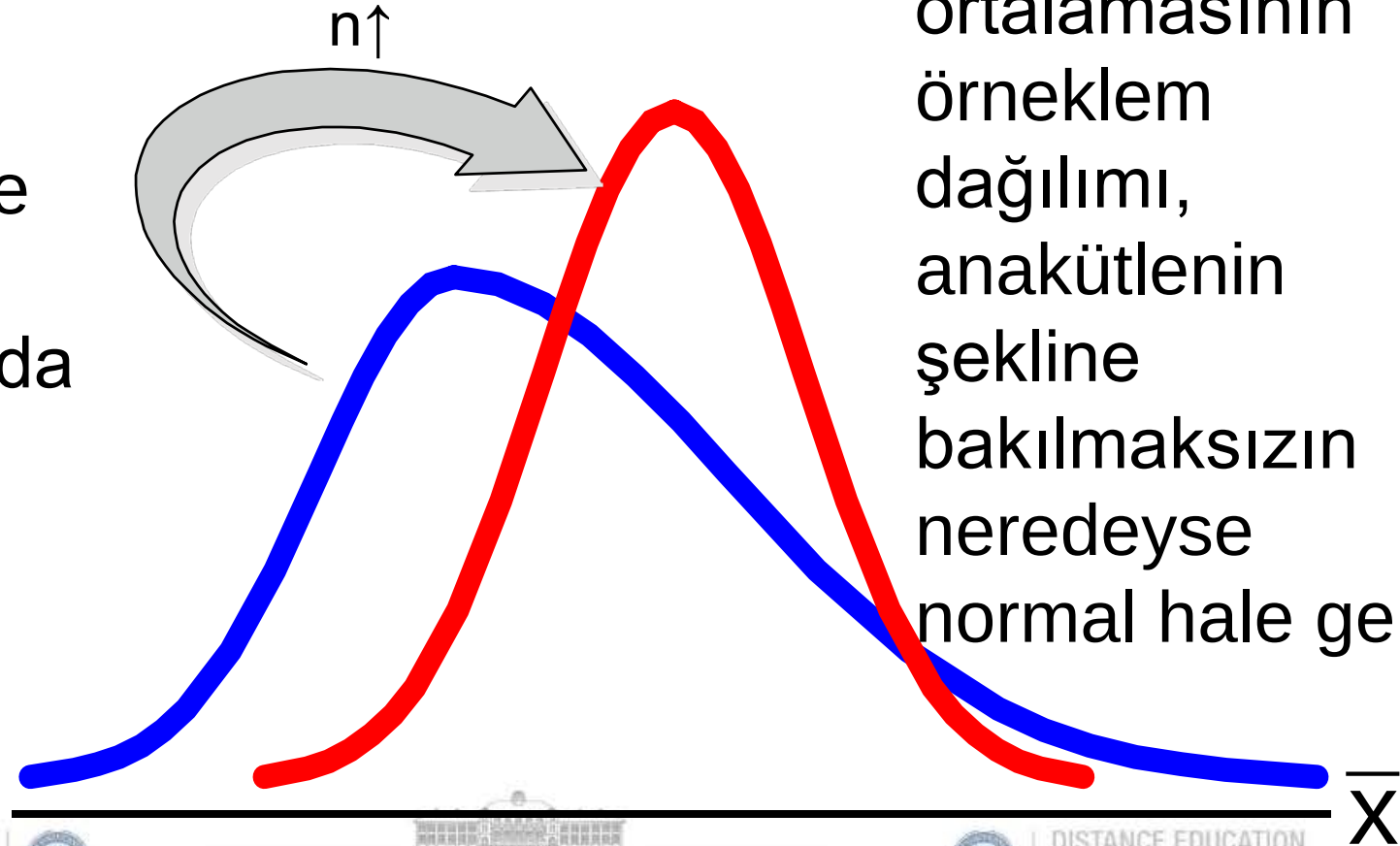
$$\mu_{\bar{x}} = \mu$$

ve

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

Merkezi limit teoremi

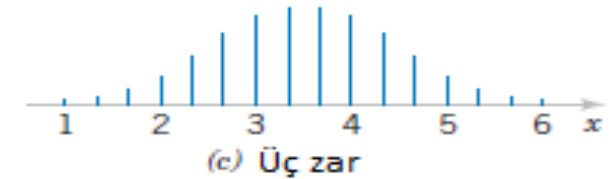
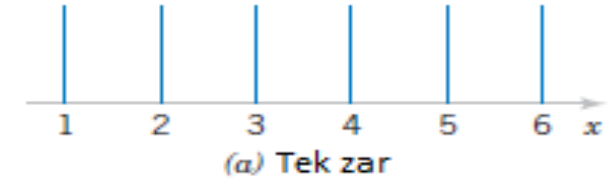
Örnek
boyutu
yeterince
büyük
olduğunda
...



örneklem
ortalamasının
örneklem
dağılımı,
anakütlenin
şekline
bakılmaksızın
neredeyse
normal hale gelir

Merkezi limit teoremi

- $\bar{X}$ için normal yaklaşım örneklem boyutu n 'ye bağlıdır.
- Aşağıdaki şekil (a) altı yüzlü hilesiz bir zarda tek atışlar için elde edilen dağılımı göstermektedir. Elde edilen tüm 1, 2, 3, 4, 5, 6 değerleri için tüm olasılık değerleri ($1/6$) eşittir.
- Şekil (b) iki zar atıldığında gelen sayıların ortalamasının dağılımını göstermektedir.
- Şekil (c), (d) ve (e) sırasıyla 3, 5 ve 10 zar atıldığında gelen sayıların ortalamasına ilişkin dağılımları göstermektedir.
- Çoğu durumda $n \geq 30$ ise dağılımın şeklini önemsemeksizin normal yaklaşım tatmin edici sonuçlar vermektedir.
- Eğer $n < 30$ ise merkezi limit teoremi popülasyonun dağılımı normal dağılımdan çok farklı değilse iyi sonuçlar vermektedir.



Örnek ortalaması örneklem dağılımı: Popülasyon normal dağılıma uymadığında

Örneklem Dağılımı
Özellikleri:

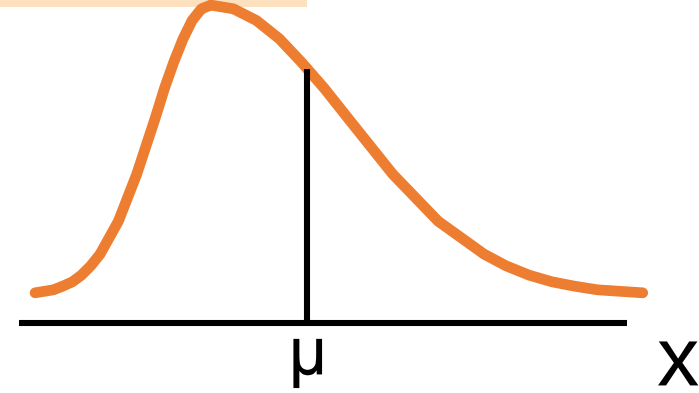
Merkezi Eğilim

$$\mu_{\bar{x}} = \mu$$

Değişim

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

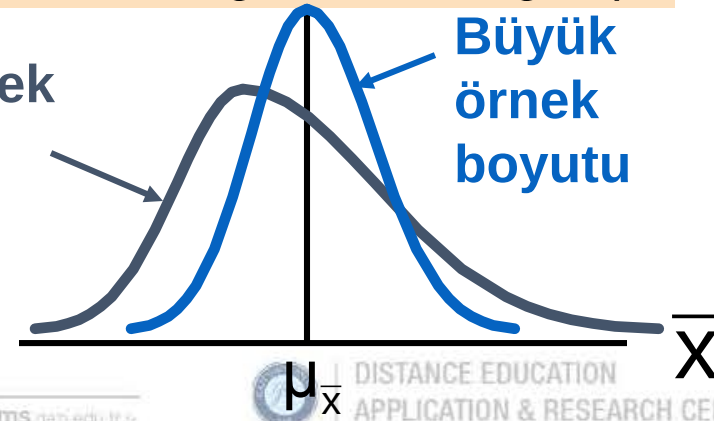
Popülasyon Dağılımı



Örneklem Dağılımı
(n büyüdükçe normal dağılım haline gelir)

Küçük örnek
boyutu

Büyük
örnek
boyutu



Örneklem ne kadar geniş olursa yeterlidir?

- Çoğu dağılıma göre, $n > 30$ neredeyse normal olan bir örnekleme dağılımı verecektir
- Yeterli derecede simetrik dağılımlar için, $n > 15$
- Normal popülasyon dağılımları için, ortalamanın örneklem dağılımı daima normal dağılım gösterir

Örnek

- Bir popülasyonun $\mu = 8$ ortalamaya ve $\sigma = 3$ standart sapmaya sahip olduğunu varsayalım. Rassal olarak $n = 36$ boyutunda bir örnek seçilmiş olsun.
- Örnek ortalamasının 7.8 ile 8.2 değerleri arasında olma olasılığı nedir?

Örnek

Çözüm:

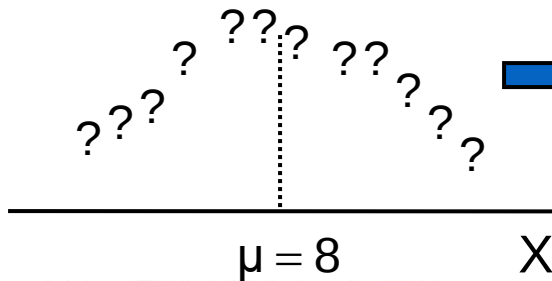
- Popülasyon normal dağılıma uymuyor bile olsa merkezi limit teoremi kullanılabilir ($n > 30$)
- ... bundan dolayı $\bar{X}$ 'in örneklem dağılımı yaklaşık olarak normal dağılmaktadır.
- ... $\mu_{\bar{X}} = 8$ ortalaması ile
- ...ve $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{3}{\sqrt{36}} = 0.5$ standart sapması ile

Örnek

Çözüm (devamı):

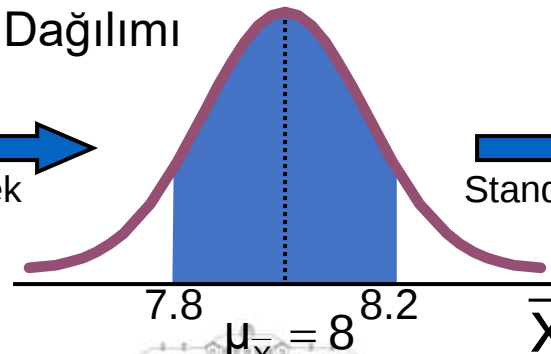
$$P(7.8 < \bar{X} < 8.2) = P\left(\frac{7.8 - 8}{3/\sqrt{36}} < \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} < \frac{8.2 - 8}{3/\sqrt{36}}\right)$$
$$= P(-0.4 < Z < 0.4) = 0.6554 - 0.3446 = \boxed{0.3108}$$

Popülasyon
Dağılımı



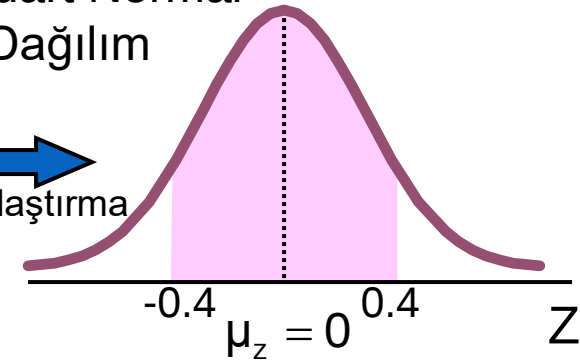
Örneklem
Dağılımı

Örnek



Standart Normal
Dağılım

Standartlaştırma



Örnek

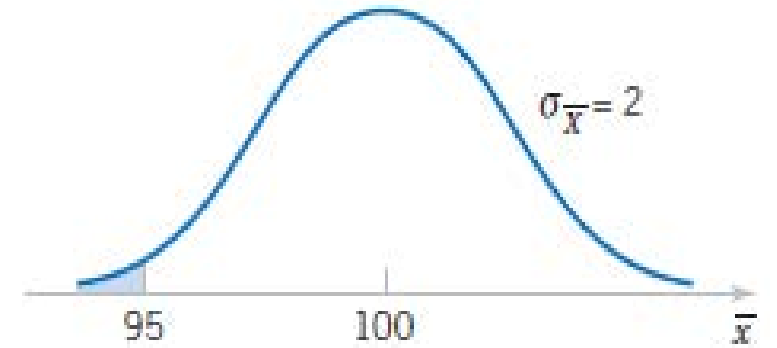
Elektronik parça üretimi yapan bir firma ortalaması 100Ω ve standart sapması 10Ω olan rezistanslar imal etmektedir. Rezistansın dağılımı normaldir. Örneklem hacmi $n = 25$ olan rezistansların oluşturduğu rassal örneklemin ortalamasının 95Ω değerinin altında olması ihtimali nedir?

Örneklem ortalaması $\bar{X}$ ortalaması $\mu_{\bar{X}} = 100 \Omega$ ve standart sapması

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{10}{\sqrt{25}} = 2$$

olan normal dağılımdır.

Böylece istenen olasılık değeri aşağıdaki şekildeki taralı alandır.



$\bar{X} = 95$ değeri için standardizasyon yapıldığında aşağıdaki değeri elde ederiz.

$$Z = \frac{95 - 100}{2} = -2,5$$

Ve böylece $P(\bar{X} < 95) = P(Z < -2,5) = 0,0062$ olarak bulunur.

Popülasyon oranları

π = popülasyonun belirli bir karakteristiğe sahip olanlarının oranı

- Örnek oranı (p) π 'nin bir tahminini yapmayı sağlar:

$$p = \frac{X}{n} = \frac{\text{Örnekte ilgili özelliği olan öge sayısı}}{\text{örnek boyutu}}$$

- $0 \leq p \leq 1$

- n yeterince geniş olduğunda p yaklaşık olarak normal dağılıma uyar

(Bu, örnekleme sonsuz bir popülasyondan yer değiştirme yaparak ya da yer değiştirme yapmadan gerçekleştiğini varsayarak oluşturulmaktadır.)

p'nin örneklem dağılımı

- Eğer aşağıdaki şartlar sağlanırsa normal dağılım ile yaklaşılr:

o

$$n\pi \geq 5$$

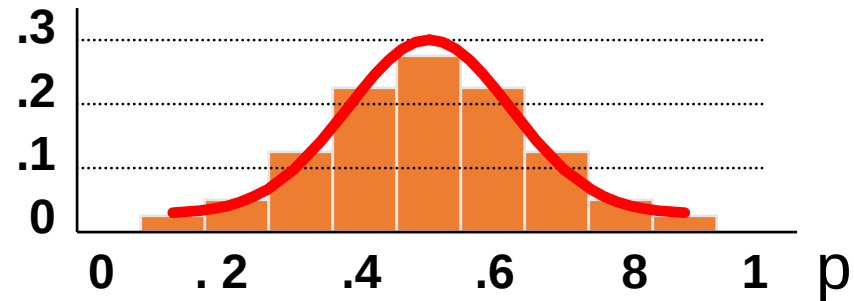
ve

$$n(1-\pi) \geq 5$$

$$\mu_p = \pi$$

$P(p_s)$

Örneklem Dağılımı



$$\sigma_p = \sqrt{\frac{\pi(1-\pi)}{n}}$$

(π = popülasyon oranı)

Oranlar için Z-değeri

p'yi bir Z değeriyle standartlaştırmak için şu formül kullanılabilir:

$$Z = \frac{p - \pi}{\sigma_p} = \frac{p - \pi}{\sqrt{\frac{\pi(1 - \pi)}{n}}}$$

Örnek

- Eğer Öneri A'yı destekleyen seçmenlerin gerçek oranı $\pi = 0.4$ ise, popülasyonun 200 birim genişlikteki bir örneğinin örneklem oranının 0.40 ile 0.45 arasında olma olasılığı ihtimali nedir?

■ yani: **eğer $\pi = 0.4$ ve $n = 200$,
 $P(0.40 \leq p \leq 0.45)$ nedir?**

Örnek

- eğer $\pi = 0.4$ ve $n = 200$,
 $P(0.40 \leq p \leq 0.45)$ nedir?

$$\sigma_p \text{ 'yi bul : } \sigma_p = \sqrt{\frac{\pi(1-\pi)}{n}} = \sqrt{\frac{0.4(1-0.4)}{200}} = 0.03464$$

Standart
Normal
dağılıma
dönüştür:

$$P(0.40 \leq p \leq 0.45) = P\left(\frac{0.40 - 0.40}{0.03464} \leq Z \leq \frac{0.45 - 0.40}{0.03464}\right) \\ = P(0 \leq Z \leq 1.44)$$

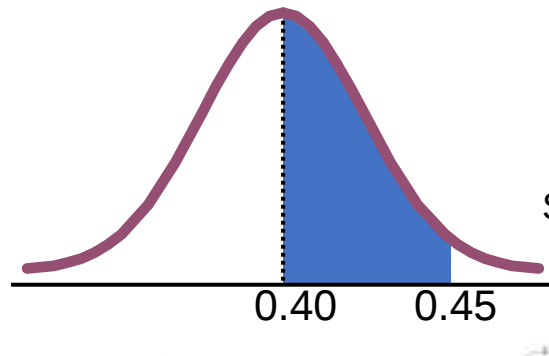
Örnek

- eğer $\pi = 0.4$ ve $n = 200$,
 $P(0.40 \leq p \leq 0.45)$ nedir?

Birikimli normal tabloyu kullanarak:

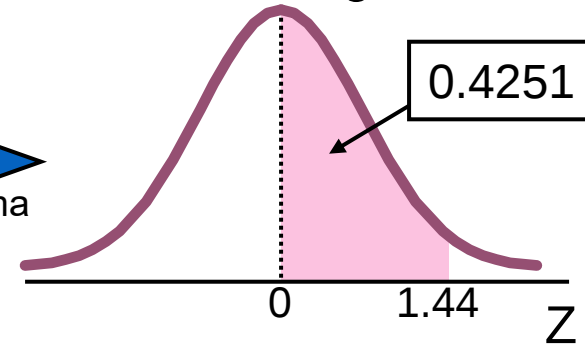
$$P(0 \leq Z \leq 1.44) = 0.9251 - 0.5000 = 0.4251$$

Örnekleme Dağılımı



Standartlaştırma

Standart Normal Dağılım



Bölüm özeti

Bu bölümde

- Örneklem dağılımlarını
- Ortalamaların örneklem dağılımlarını
 - o Normal dağılıma uyan popülasyonlar
 - o Merkezi Limit Teoreminin kullanımı
- Bir oranın örneklem dağılımını
- Örneklem dağılımını kullanarak olasılıkların hesaplanmasını inceledik



Sonlu Popölasyonlardan Örnekleme Yapmak



Öğrenim hedefleri

Bu konuda :

- Sonlu popülasyon düzeltmelerinin ne zaman gerekli olduğu
- Standart hataların hesaplanmasında sonlu popülasyon düzeltme faktörlerinin nasıl kullanılacağı

öğrenilecektir.

Sonlu popölasyon düzeltme faktörleri

- Hem örnek ortalamasının hem de örnek oranının standart hatasını hesaplamak için kullanılır
- Örneklem büyüklüğü, n , anakütle büyüklüğünün, N , % 5'inden büyük olduğunda gereklidir (yani $n / N > 0.05$)
- Sonlu Popölasyon Düzeltme Faktörü:

$$fpc = \sqrt{\frac{N - n}{N - 1}}$$

fpc'nin standart hataların hesaplanmasında kullanımı

Sonlu Popülasyonlarda Ortalamanın Standart Hatası

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$$

Sonlu Popülasyonlarda Orantıların Standart Hatası

$$\sigma_p = \sqrt{\frac{\pi(1-\pi)}{n}} \sqrt{\frac{N-n}{N-1}}$$

fpc'nin kullanımı standart hatayı düşürmektedir

- fpc her zaman 1'den düşüktür
- Böylece, kullanıldığında standart hatayı düşürmektedir
- Popülasyon parametrelerinin çok daha hassas tahminlerinin yapılması sonucunu vermektedir

fpc'nin ortalama ile kullanımı: Örnek

1000 birim büyüklüğündeki bir popülasyondan 40 standart sapma ile 100 boyutlu rassal bir örnek alındığını varsayalım.

Burada $n=100$, $N=1,000$ and $100/1,000 = 0.10 > 0.05$.

Böylece ortalamanın standart hatası için fpc'yi kullanarak:

$$\sigma_{\bar{X}} = \frac{40}{\sqrt{100}} \sqrt{\frac{1000-100}{1000-1}} = 3.8$$

Konu özeti

Bu konu içerisinde

- Bir sonlu popölasyon düzeltmesinin ne zaman kullanılması gerektiği.
- Hem bir örnekleme ortalamasının hem de bir örnekleme oranının standart hatasını hesaplarken sonlu bir popölasyon düzeltme faktöründen nasıl yararlanılacağı

tartışılmıştır.